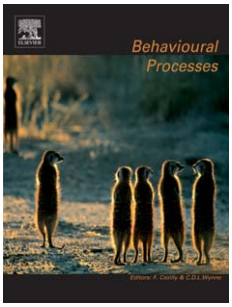

Title: Pictorial gaze cues do not enhance long tailed macaques' performance on a computerised object-location task

Authors: Sarah-Jane Vick, Ido Toxopeus, James R. Anderson

PII: S0376-6357(06)00166-5
DOI: doi:10.1016/j.beproc.2006.07.003
Reference: BEPROC 1628

To appear in: *Behavioural Processes*

Received date: 10-7-2006
Accepted date: 19-7-2006

Please cite this article as: Vick, S.-J., Toxopeus, I., Anderson, J.R., Pictorial gaze cues do not enhance long tailed macaques' performance on a computerised object-location task, *Behavioural Processes* (2006), doi:10.1016/j.beproc.2006.07.003

Pictorial gaze cues do not enhance long tailed macaques' performance on a computerised object-location task

Sarah-Jane Vick†* Ido Toxopeus‡, James R. Anderson†

† Department of Psychology, University of Stirling, Stirling, FK9 4LA Scotland

‡ Ethology Station, University of Utrecht, Munsterlaan 19, 3584 CP, Utrecht, The Netherlands

* Corresponding author: E-mail: sarah-jane.vick@stir.ac.uk

Telephone: ++44 (0)1786 467 648

Fax: ++44 (0)1786 467 641

Abstract

The perception of pictorial gaze cues was examined in long-tailed macaques (Macaca fascicularis). A computerised object location task was used to explore whether the monkeys would show faster response time to locate a target when its appearance was preceded with congruent as opposed to incongruent gaze cues. Despite existing evidence that macaques preferentially attend to the eyes in facial images and also visually orient with depicted gaze cues, the monkeys did not show faster response times on congruent trials either in response to schematic or photographic stimuli. These findings coincide with those reported for baboons testing with a similar paradigm in which gaze cues preceded a target identification task (Fagot and Deruelle 2002). When tested with either pictorial stimuli or interactants, non human primates readily follow gaze but do not seem to use this mechanism to identify a target object; there seems to be some mismatch in performance between attentional changes and manual responses to gaze cues on ostensibly similar tasks.

Keywords: nonhuman primate, gaze following, object-choice task, cueing.

Introduction

The study of gaze understanding in nonhuman primates has been assessed using two basic methodologies: a gaze following paradigm, in which the respondent visually co-orients with a model's gaze (Itakura, 1996), and various object-choice tasks which require the animal to select an object on the basis of gaze cues (Anderson et al. 1995). These two approaches have revealed what has been described as a 'functional dissociation' (Gomez 2005); primates readily gaze follow of both humans and conspecifics (Tomasello et al, 1998) but successful performance on object-choice tasks is less consistent and may require learning over a number of sessions (e.g. Vick and Anderson 2000)[1]. There is apparently some divergence between knowledge and action as measured by the dependent variables of visual attention and manual responses, respectively.

Several species of nonhuman primates have been shown to change their visual orientation to track a human interactant's head and eye direction or eye gaze alone (Ferrari et al. 2000; Okamoto et al. 2002; Tomesello et al, 1999). Similar studies which have used live conspecific models are largely unable to control which cues are present or attended to so that it is unclear which factors (such as bodily orientation and head direction) influence responses (Tomasello et al. 1998). As conspecific cues are difficult to vary systematically, studies have also used photographic images rather than interactants and have reported clear evidence for visual co-orientation (Deaner and Platt 2003; Emery et al. 1997; Lorincz et al. 1999; Scerif et al. 2004; Shepherd, Deaner and Platt, 2006). In addition, presenting video or photographic stimuli may allow more control over timing and cues, while removing any social constraints upon interactions with either humans or conspecifics (Fagot and Deruelle 2002). Although there is still debate regarding the understanding of pictorial representations in nonhumans (Bovet and Vauclair 2000; Martin-Malivel and Fagot 2001; Pascalis et al. 1999), it is clear that

both video and static images are valuable in the study of social cognition in monkeys (Capitanio 2002; Dasser 1987; Kyes et al. 1992).

There are consistent patterns in terms of how monkeys look at images of faces; macaques reliably allocate most attention to the eyes, even when images were inverted or scrambled (Guo et al. 2003). Moreover, this same basic pattern is found, in terms of viewing time, number of fixations, and other temporal and spatial characteristics, whether the facial stimuli presented is conspecific, chimpanzee, human or even simple schematic images (Keating and Keating 1982; Gothard et al. 2004; Guo et al. 2003; Sato & Nakamura, 2001). Human facial images have sometimes been used in gaze studies as cues may be more readily controlled (Fagot and Deruelle 2002; Vick et al 2001). To date, very few studies using conspecific stimuli have manipulated both head and eye direction; using scanning patterns as a measure, rhesus macaques have been shown to respond to both head and eye direction alone by shifting attention to a congruent quadrant of space (Lorincz et al 1999). Deaner and Platt (2003) also found that non-predictive head and eye-gaze cues enhanced gaze shifts to detect peripheral targets by rhesus monkeys (and humans tested within the same paradigm).

While monkeys readily co-orient and visually inspect cued locations, the attention to the target object itself appears ambiguous at best. Emery et al (1997) demonstrated that objects which appeared in a location congruent with depicted visual attention were inspected more than identical distractors. However, most attention was directed at central monkey image prior to object appearance and there was no significant difference between target and distractor when only these remained on screen. Similar results with Diana monkeys also show coorientation to a spatial location congruent with the gaze direction of a photographed conspecific, in terms of significantly more first looks, and more frequent and longer visual inspections. However, attention to target and anti-target locations *following* the appearance of the object itself was not

[1] While success is more readily found with competitive paradigms that require non-attended items to be selected (Flombaum and Santos, 2005; Hare et al, 2000; 2001; Vick and Anderson, 2003), the nature of the relationship

reported; thus, it impossible to ascertain the relationship between gaze following and subsequent levels of interest in the target rather than just the cued location (Scerif et al. 2004).

Measurement of eye movement and other forms of visual co-orientation can certainly inform about nonhuman primates' responses to gaze as a social signal, but the absence of any subsequent response to targets means that it is difficult to assess how attention shifts are translated into action. Moreover, the timing of both independent and dependent variables differ in gaze following and object choice approaches. Object choice studies typically involve presenting a cue for a few seconds to ensure that the subject has time to attend, but even relatively brief delays between cue and response is detrimental to chimpanzee performance (Barth et al. 2004). In comparison, studies using visual fixations as a dependent measure have shown effects within a few seconds and even at first fixation (e.g. Deaner and Platt 2003; Scerif et al. 2004). It may be that temporal contiguity of target and cue is crucial as attention shifts occur rapidly and then quickly dissipate.

One way to examine whether manual responses or timing are detrimental to object-choice performance is to modify procedures in order to make the gaze following and object response tasks more comparable. Only one study to date has combined pictorial gaze stimuli with a manual response, rather than measuring visual scan patterns alone. The impact of schematic and photographic (human) gaze cues upon accuracy and response times was examined in baboons using a joystick to respond on a go no-go target identification task (Fagot and Deruelle, 2002). This approach resembles the object choice task in terms of gaze cues preceding a manual response, but also allows more subtle differential effects of cue type on reaction times (RTs) to be measured. Moreover, the presentation of gaze cues was much briefer than in traditional manual response tasks with the delay between the appearance of cue and

between competitive tasks and a tendency to co-orient with others is less clear.

target, or stimulus onset asynchrony (SOA), being 300ms.[2] However, Fagot and Deruelle (2002) found that targets were not detected more accurately or more rapidly when preceded with congruent as opposed to incongruent gaze cues.

The current study aims to further explore this type of cueing paradigm by examining the relationship between gaze cues and manual responses in long-tailed macaques. Firstly, the only cue presented by Fagot and Deruelle was eye gaze, but a fuller exploration considering a variety of attention cues and their interactions is warranted (Langton 2000; Lorincz et al. 1999; Perrett and Emery 1994). The present study therefore presents a variety of head and eye direction cues to examine the relative influence of these information sources. Schematic faces and matched scrambled control images were also presented to examine the basic features that might affect responses, such as face-like configuration and overall perceptual asymmetry of stimuli (Quadflieg, Mason & Macrae 2004). Secondly, the current study uses a wider ranger of SOAs to explore more systematically the issue of timing of cues and responses (drawing upon human literature on reflexive gaze following, see also Deaner and Platt 2003). Thirdly, whereas the baboons were tested within a pre-learned go-no go joystick task which required object discrimination (Fagot and Deruelle 2002); we used a less cognitively demanding and more strictly spatially based task requiring a response directly to the target location using a touch-screen. If target location is facilitated by co-orientation we should see faster reaction times to congruent than incongruent targets, and this pattern may vary systematically with differing cue-target timings (Deaner and Platt 2003).

Materials and Methods

Subjects

[2]SOA later modified to 800ms, the cue-target distance was also reduced and the baboons were trained to attend to the central area of the monitor but these changes had no impact on the results obtained.

The monkeys tested were 9 long-tailed macaques (Macaca fascicularis: 8 females and 1 male, age range 9-26 years, mean age 14.6 years) housed in a large social group (46 individuals) at the Ethology Station, University of Utrecht, The Netherlands. The group had access to a large indoor/outdoor enclosure ($80m^2$) but could be isolated in a smaller indoor experimental area (0.5m x 1.2m) for testing, while maintaining auditory and limited visual contact with their group. The monkeys were fed twice daily with commercial monkey pellets and twice a week this was supplemented with fruit or vegetables and grain. Treats such as nuts, raisins and cereal were used as reinforcers during testing.

Apparatus

A 14-inch (640 x 480 pixels) Apple Macintosh monitor was secured in front of an opening in the mesh of the test area. The monitor was fitted with a Micro-Touch touch-screen and was connected to an Apple Macintosh computer which controlled presentation of trials and recorded responses using Authorware Professional software. Correct responses were rewarded with food items (grain and raisins) delivered by a Universal Feeder (G5912 Gebrands Corporation) that was connected to the computer. A video camera was secured above the test area so that the experimenter could monitor the monkeys' general behaviours as they performed the task.

Procedure

Training

The monkeys had previously been trained and tested on delayed non match to sample tasks using the touch-screen (Veenema, 1998). The monkeys were re-familiarised firstly to the experimental area and subsequently to the apparatus. Once the monkeys reliably responded to the touch-screen and could be readily separated from the group for at least 10 minutes, testing commenced.

Testing

Once a monkey had entered the test area, they were separated from the rest of the group using a sliding partition. The computer program was then initiated and the experimenter observed the session from an adjacent room via the video link. The monkey remained in the test area until a maximum of 50 trials had been presented or 25 minutes had elapsed. If the monkey ceased responding for over 3 minutes or was became distressed (for example, when group fights broke out in the adjacent enclosure), the session was terminated and the monkey was released back into the group. As entry into the testing area was voluntary, not all animals were tested on all days (data from additional subjects completing only a few sessions each were excluded from analysis). Subjects were tested with a schematic stimulus set for 5 weeks and with photographic images for the following 4 weeks; order of presentation was not counterbalanced due to the group testing situation.

Stimuli

Schematic faces: these were line drawings (150 x 150 pixels) of faces and scrambled faces. Pupils appeared positioned in one of five locations: up, down, left, right or central (a neutral condition which offered no directional cue towards peripheral locations) relative to the centre point of the face (depicted by the nose). Photographs: these were colour images (150 x 150 pixels) of a human female with variations in head and eye direction. In this set, the face cued only along the horizontal axis, e.g. head and/or eyes to the right/left. Any asymmetry in the photographs was accommodated by using reversed right images to create left images (see Scerif et al 2004). There were three variations of cue available: congruent head and eye cue with the head turned in full profile to left or right, eyes only directed to either the left or right or displaying direct gaze, and half-profile images in which the head was turned 30 degrees from centre with eyes either congruent with head turn or fixated centrally (direct gaze).

The task

Following a 30-second delay during which the monitor was uniformly black, a schematic face or control (scrambled face) stimulus appeared; the monkey touching this central stimulus initiated the trial. After a 500msec delay, pupils appeared in the face or control image. Following another delay of 100ms, 300ms or 1000ms (stimulus onset asynchrony: SOA) the central image was removed and a target square (100 x 100 pixels) appeared in one of the four locations (up, down, left, right). Figure 1 shows an example of a presentation sequence. For the photograph stimulus set the pre-trial stimulus was modified; a white square that appeared in the centre of the screen was used. The square disappeared when touched and the trial commenced with the stimulus photograph appearing in the centre of the screen after a 500ms delay (see Figure 2).

Each cue type for both stimulus sets was presented as congruent and incongruent (i.e. cues were non-predictive) with the subsequent target at each of the three SOAs. For the schematic images, the trials varied according to four factors: Stimulus type (Face and Control), timing of SOA (100ms, 300ms and 1000ms), and Cue condition (Congruent, Incongruent and Neutral); the four target locations (up, down, left, right) were used to ensure that cues were non-predictive and were not considered in the subsequent analyses. For the photographic stimuli, the trials varied in terms of face type (profile, half profile, frontal), timing of SOA (100ms, 300ms and 1000ms), and Cue condition (Congruent, Incongruent and Neutral). The monkey responded by touching the target, which then disappeared, and a reward was then delivered by the automatic feeder. The next trial commenced after a 30-second inter-trial interval. If the monkey failed to touch the initial image within 30 seconds, it re-appeared after a 10-second time-out. If the monkey initiated a trial but subsequently failed to touch the target within 30 seconds a new trial was presented without any reward being delivered. Trials were

pseudo-randomly presented with no more than three consecutive trials with the same stimulus type, SOA, target location or congruence value. Responses on each trial were recorded in terms of latency and the co-ordinates of each touch to the screen; response time (RT) was measured in ms and timed from the target onset.

Fig. 1 and Fig 2. about here

Data Analysis

Target location response times of less than 100msec (anticipations) or greater than 3000msec (time outs) were excluded from the analysis; only rapid responses which were likely to be related to gaze cues were considered, with a 3 second interval comparable to the duration used in scan measurement studies. The percentage of trials eliminated by this procedure was 11.6% (347/3003) and 7% (226/3226) for schematic and photographic stimuli, respectively. Incorrect responses (first touches to areas of the screen other than to the target area) were also excluded from analysis, removing a further 9.4% (281) trials and 11.4% (368) trials (with only 16 and 14 trials being touches to locations cued by incongruent gaze cues) for schematic and photographic conditions, respectively.

Mean RTs were calculated for each individual in each condition; means were used instead of medians as trial numbers were unequal (Miller, 1998). Repeated measures analyses of variance were conducted on the mean RT data. Repeated measure Anovas were also conducted on error rates (Friesen and Kingstone, 1998) for schematic (all $p>0.1$) and photographic images (all $p> 0.05$) and revealed no systematic variation in error rates. The data was also analysed for each individual monkey; there were a small number of significant

interactions but no response patterns indicative of reflexive gaze and individual performances are therefore not reported further here.

Results

1. Schematic images

A repeated measures anova conducted on mean RTs (See Table 1) revealed no significant main effects, that is, response times did not differ according to whether or not a face or control image was presented ($F_{1,8} = 0.01$, $p = 0.98$), SOA ($F_{2,16} = 0.77$, $p = 0.48$) or as a function of the cue's predictive value ($F_{2,16} = 2.77$, $p = 0.09$). There were no significant interactions (Face x SOA, $F_{2,16} = 1.52$, $p = 0.25$, Face x Cue, $F_{2,16} = 1.83$, $p = 0.19$, Cue x SOA, $F_{4,32} = 0.88$, $p = 0.48$, Face x Cue x SOA, $F_{4,32} = 1.12$, $p = 0.37$). There was no evidence that responses differed to schematic and scrambled images. Further, there is no evidence to suggest that congruence between cue and target in terms of either direction or timing had any impact upon response times in this target location task.

Table 1

---

2. Photographic images

Each face type was examined separately in order to allow all cue combinations to be thoroughly analysed.

Head direction: As head direction was either congruent or incongruent with the target location, a 3 (SOA) x 2 (Cue) repeated measures Anova was conducted on group data (see Table 2 for mean RTs). There were no significant main effects; RTs did not differ according to SOA ($F_{2,16} = 3.12$, $p = 0.086$) or whether the target was congruent or incongruent ($F_{1,8} = 0.32$, $p = 0.59$). There was also no significant interaction between SOA and Cue condition ($F_{2,16} =$

2.37, p = 0.13). These results suggest that depicted head direction did not impact upon response times, in terms of either spatial or temporal cue-stimulus contiguity in the target location task presented.

Eye direction: As eye direction could either be congruent, incongruent or neutral regarding the target location, a 3 (SOA) x 3 (Cue) repeated measures Anova was conducted on group data (see Table 2 for mean RTs). Exploration of the SOA main effect showed that 100ms SOA trials were responded to faster than longer SOAs (mean 100ms = 938ms, 300ms = 1011ms and 1000ms = 1008) but this effect did not reach significance ($F_{2,16}$ = 3.40, p = 0.059). There was a main effect of Cue condition, with neutral stimuli receiving the fastest responses (Neutral = 946ms, Congruent = 999ms and Incongruent = 1012ms: $F_{2,16}$ = 3.86, p = 0.043). Fisher's tests revealed that responses were faster to neutral compared with incongruent targets (t_1 = 4.16, p = 0.03), while responses to congruent targets were not significantly different from either. Overall, the data from the eye-direction conditions fail to demonstrate that congruent cueing of the target led to faster response times, although it does seem that neutral images (direct gaze) may have speeded responses.

Half-profile: As head and eye direction could be either congruent or incongruent with one another, the data were analysed on the basis of head orientation alone (i.e. ignoring eye direction), and also taking account of whether or not eye gaze was compatible with head direction. When the data were analysed on the basis of head direction alone in a 3 (SOA) x 2 (cue condition) repeated measures Anova, there were no significant effects or interactions (SOA $F_{2,16}$ = 1.54, p = 0.24; Cue $F_{1,8}$ = 0.73, p = 0.42; SOA x Cue $F_{2,16}$ = 2.13, p = 0.15; see Table 2 for mean RTs). To examine whether head and eye direction interacted, a 3 (SOA) x 4 (Cue condition: head and eye directions) repeated measures Anova was also conducted; there

were no significant main effects or interactions RTs did not differ according to SOA or Cue (SOA $F_{2,16} = 1.203$, $p = 0.33$; Cue $F_{3,24} = 0.84$, $p = 0.49$; SOA x Cue $F_{6,48} = 1.02$, $p = 0.42$). Unlike direct gaze within a frontal face, direct gaze as depicted within a half-turned face did not influence RTs, even when data were analysed simply on the basis of whether the eyes were engaging in direct gaze or were averted ($F_{1,8} = 1.51$, $p = 0.254$). These results suggest that half-profile cues do not facilitate faster responses, and further, that systematically varying the eye direction with the half-profile face does not influence response patterns in this target location task.

Table 2

Discussion

The current study aimed to tap into spontaneous gaze following responses and examine how these are translated into subsequent actions on attended objects. The results with both schematic and photographic stimuli fail to demonstrate any advantage of congruent gaze cues upon the macaques' subsequent response times to locate a target. Responses were slightly faster overall for photographic rather than schematic images, however due to possible order effects it is difficult to conclude whether this is a meaningful difference in terms of face processing. These negative findings are in line with previous research using a target identification task in baboons (Fagot and Deruelle 2002), suggesting that our spatially based target location task did not differ substantially from their identification task. In addition, the macaques were no more likely to respond to the more salient cue of profile views than eye gaze alone, and there was no systematic influence of temporal contiguity of cue and target. That is, these modifications did not impact upon response times following congruent and incongruent cues.

A potential problem with the current study is that the current task did not require the central gaze stimuli to be actively processed (due to brief SOA), as responses were only needed to start a trial and to locate the target. Firstly, this seems unlikely as the monkeys were monitored via a video link throughout the session; they were highly motivated to perform the task and, while eye direction was not measured directly, the monkeys were reliably oriented towards the screen for the vast majority of a session. It could also be argued that a gaze following response could lead the monkeys to look beyond the computer screen rather than to the target object per se. While eye measurement studies can control for this by training subjects to only foveate within defined boundaries (Gothard et al 2004), the current study was dependent upon natural gaze shifts. In any case, it is likely that gaze following should still lead to differential response times in detecting cued and uncued targets; orienting in the wrong direction would slow detection more than orienting in the same general direction as the target. Secondly, in humans reflexive gaze is robust even when participants are specifically asked to ignore the central image. As identified in the introduction, primates have an innate interest in viewing social stimuli such as faces, with the eyes consistently of greatest interest (e.g. Gothard et al. 2004). The finding that direct gaze photographs did shorten response times further suggests that the monkeys were processing the central stimuli on some level; direct gaze is considered a distinct percept which captures attention (Baron-Cohen 1994; von Grunau & Anston, 1995).

Together, these studies with baboons and macaques seem to conflict with growing evidence that monkeys visually co-orient in response to pictorial stimuli. Although Fagot and Deruelle (2002) suggested that co-orientation in other studies may reflect learning, this seems unlikely as co-orientation responses are not subject to differential reinforcement (Emery et al. 1997; Lorincz et al. 1999; Scerif et al. 2004). Both the current study and Fagot and Deruelle's (2002) also avoided differential reinforcement by using non-predictive cues and rewarding both

cued and uncued target location and identification, respectively. However, the baboons were subsequently shown to learn to use these same gaze cues when these were repeatedly presented contingently with the target (Fagot and Deruelle 2002), suggesting that they can learn to use gaze in a problem solving situation (Vick & Anderson, 2000).

Two factors which have received attention in the gaze following literature appear particularly relevant for our findings: whether the problem is presented in a competitive or cooperative context and whether the model is human or conspecific. In relation to recent research using more competitive paradigms, and reporting better performance when responses are made to an unattended object (e.g. Hare, Call and Tomasello 2000; Flobaum and Santos 2005; Vick and Anderson 2003), it could be argued that in fact slower responses should be expected to cued objects. If there is some inhibition of responses to attended objects, avoiding the direction of another's gaze could potentially lead to faster detection of the target on incongruent trials. However, this is an unlikely explanation for the absence of cueing benefits due to the robust nature of the visual co-orienting response in primates, and the absence of any systematic variation in favour of uncued targets either.

Perhaps the inter-specific stimuli deserves further attention; the visible morphology of the human eye is distinct from those of other primates, notably in the amount of visible white sclera (Kobayashi and Koshima 2001; Perrett and Mistlin 1990), and this may mean that the psychophysics of gaze detection will similarly differ (e.g. see Ricciardelli et al. 2000 for human gaze perception). Nonetheless, there is evidence that macaques scan human and monkey faces is a comparable manner (Guo et al. 2003) and moreover respond to a wide range of gaze stimuli with visual co-orientation (Lorincz et al. in press). While the social context of gaze in certainly due further study (e.g. Shepherd, Deaner & Platt, 2006), we suggest instead that the considerable variations in performance across methodologies are likely to result from the different response demands of visual co-orientation studies and object based paradigms. While

it is difficult to understand how space-based responses would not improve target detection within that space, this dichotomy has even been shown within the same subjects (Povinelli et al. 1999, see also Call et al. 1998). Moreover, both the current study and Fagot and Deruelle's (2002) indicate that even when object response tasks have a temporal framework which is comparable to that of eye movement studies, this central difference remains.

An alternative approach to examining gaze following and subsequent actions is to ask whether nonhuman primates are able to recognise the relationship between gaze and targets in other individuals. For example, using an expectancy violation paradigm in which an individual is shown attending to one of two objects before manipulating either the attended or ignored object. During familiarisation trials, only attended objects are acted upon so that the viewer's visual interest diminishes; understanding the relationship between attention and action should lead to increased interest (measured in looking patterns) when incongruent gaze-action trials occur in which the previously ignored object is manipulated. However, findings using this paradigm are inconsistent; cotton-top tamarins have shown differential responses (Santos and Hauser 1999); but capuchins and squirrel monkeys do not (Anderson et al. 2004). As with other co-orientation responses, it is not clear whether it is unusual changes in orientation within trials that are unexpected rather than an understanding between gaze and object of attention per se. Interestingly, Scerif et al. (2004) also report something akin to expectancy violation in Diana monkeys with the monkeys checking back (see also Call et al. 1998) to the stimulus monkey photograph when a target appeared at a incongruent location. This indicates that the relationship between attention and target is processed in some manner, but further study is needed to ascertain precisely what this might mean in terms of cognitive processing. Thus, research focusing upon the recognition rather than production of the gaze-action relationship has also failed to reveal a clear pattern. Nonhuman primates are clearly sensitive to others' direction of gaze but the underlying mechanisms remain unclear. In order to better understand

the functional dissociation between perception and action in response to gaze information, future research should attempt to measure these responses in parallel.

Acknowledgements

This study was conducted at the Ethology Station at the University of Utrecht, The Netherlands. We would like to thank J.A.R.A.M. van Hooff and all the staff and students for their assistance with this study. This study complies with current laws on animal research in The Netherlands.

References

Anderson JR, Mitchell RW (1999). Macaques but not lemurs co-orient visually with humans. Folia Primat. 70: 17-22.

Anderson JR, Sallaberry P, Barbier H (1995) Use of experimenter-given cues during object-choice tasks by capuchin monkeys. Anim Behav 49:201-8

Anderson JR, Kuroshima H, Kuwahata H, & Fujita K (2004). Do squirrel monkeys (*Saimiri sciureus*) and capuchin monkeys (*Cebus apella*) predict that looking leads to touching? *Anim Cogn*, 7, 185 – 192.

Baron-Cohen S (1994). How to build a baby that can read minds: cognitive mechanisms in mindreading. Curr Psych of Cogn 13: 513-552.

Barth J, Reaux JE, Povinelli DJ (2004) Chimpanzees' (*Pan troglodytes*) use of gaze cues in object-choice tasks: different methods yield different results. Anim Cogn 8: 84-92

Bovet D, Vauclair J (2000). Picture recogntion in animals and humans. Behav Brain Res 109: 143-165.

Call J, Hare BA, Tomasello M (1998) Chimpanzee gaze following in an object-choice task. Anim Cog 1: 89-99.

Capitanio JP (2002). Sociability and responses to video playbacks in adult male rhesus monkeys (*Macaca mulatta*). Primates 43: 169-177.

Dasser V (1987). Slides of group members as representations of the real animals (*Macaca fascicularis*). Ethol 76: 65-73.

Deaner RO, Platt, ML (2003) Reflexive social attention in monkeys and humans. Curr Biol 13: 1609-1613.

Emery N (2000). The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci and Behav Rev 24: 581-604.

Emery NJ, Lorincz EN, Perret DI, Oram MW, Baker CI (1997). Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). J of Comp Psych 111: 286-293.

Fagot J., Deruelle C. (2002). Perception of pictorial gaze by baboons (*Papio papio*). J of Exp Psych 28: 298-308.

Fagot J, Martin-Malivel J, Dépy D (1999). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? Curr Psych of Cogn 18: 923-949.

Flombaum JI, Santos LR (2005) Rhesus Monkeys Attribute Perceptions to Others. Curr Biol 15: 447-452.

Friesen CK, Kingstone A (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychonom Bull and Rev 5: 490-495.

Gómez J-C (2005) Species comparative studies and cognitive development. Trends in Cogn. Sci 9:118-125.

Gothard KM, Erickson CA, Amaral DG (2004) How do rhesus monkeys (*Macaca mulatto*) scan faces in a visual paired comparison task? Anim Cogn 7: 25-23

von Grunau M, Anston C (1995). The detection of gaze direction: A stare in the crowd effect. *Perception*, 24, 1297-1313.

Guo K, Robertson RG, Mahmoodi S, Tadmor Y, Young MP (2003). How do monkeys view faces? A study of eye movements. Exp Brain Res 150: 363-374.

Hare B, Call J, Agnetta B, Tomasello M (2000). Chimpanzees know what conspecifics do and do not see. *Anim Beh*, 59, 771-785.

Hare B, Call J, Tomasello M (2001). Do chimpanzees know what conspecifics know? Anim Beh 61: 139-151.

Kobayashi H, Kohshima S (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies of external morphology of the primate eye. *J Human Evol*, 40, 419-435.

Kyes RC, Mayer EK, Bunnell BN (1992). Perception of stimuli presented as photographic slides in cynomolgus macaques (*Macaca fascicularis*). Primates 33: 407-412.

Langton SRH (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. Quart. J Exp. Psych. 53A: 825-845.

Langton SRH, Bruce V (1999). Reflexive visual orienting in response to the social attention of others. Visual Cogn. 6: 541-567.

Lorincz EN, Baker CI, Perret DI (1999). Visual cues for attention following in rhesus monkeys. Curr Psych of Cogn 18: 973-1003.

Lorincz EN, Jellema T, Gomez J-C, Barraclough N, Xiao D, Perret, DI (in press) Do monkeys understand actions and minds of others? Studies of single cells and eye movements. In From Monkey Brain to Human Brain (Dehaene S, ed.), Oxford University Press.

Martin-Malivel J, Fagot J (2001) Perception of pictorial human faces by baboons: Effects of stimulus orientation on discrimination performance. Anim Learn Behav 29: 10-20

Miller JO (1988). A warning about median reaction time. J Exp. Psych: Human Perception and Performance, 14, 777-778.

Okamoto S, Tomonaga M, Ishii K, Kawai N, Tanaka M, Matsuzawa T (2002). An infant chimpanzee (*Pan troglodytes)* follows human gaze. Anim. Cogn. 5: 107-114.

Pascalis O, Petit O, Kim J, Campbell R (1999). Picture perception in primates: the case of face perception. *Cahiers de Psychologie Cognitive* 18, 889-922

Perrett DI, Emery NJ (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. Curr Psych Cogn, 683-694.

Perrett DI, Mistlin, AJ (1990). Perception of facial characteristics by monkeys. In WC Stebbins and MA Berkley (Eds.) Comparative Perception - Volume II: Complex Signals . (pp. 187-213). New York: John Wiley and Sons.

Povinelli DJ, Eddy TJ (1996). What Young Chimpanzees Know about Seeing. Chicago: University of Chicago Press.

Povinelli DJ, Bierschwale DT, Cech CG (1999) Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. British J Dev Psych. 17: 37-60.

Quadflieg S, Mason MF, Macrae CN (2004) The owl and the pussycat: Gaze cues and visuospatial orienting. Psych Bull & Rev, 11, 826-831

Ricciardelli P, Baylis G, Driver J (2000). The positive and negative of human expertise in gaze perception. Cogn. 77: B1-14.

Santos LR, Hauser MD (1999) How monkeys see the eyes: Cotton-top tamarins' reaction to changes in visual attention and action. Anim Cogn, *2*, 131 - 139.

Scerif G, Gomez J-C, Byrne RW (2004). What do Diana monkeys know about the focus of attention of a conspecific? Anim Behav 68: 1239-1247

Shepherd SV, Deaner RO and Platt ML (2006) Social status gates social attention in monkeys. Curr Biol 21, 19-20.

Tomasello M, Call J, Hare B (1998). Five primate species follow the visual gaze of conspecifics. Anim Beh 55: 1063-1069.

Veenema H (1998). Social influences of aging in group-living long-tailed macaques. Doctoral thesis, Universiteit Utrecht.

Vick S-J, Anderson JR (2000). Learning and limitations of eye gaze an an experimenter-given cue in an object choice task by capuchin monkeys (*Cebus apella*). J Comp Psych 114: 200-207.

Vick S-J, Bovet D, Anderson JR (2001). Discrimination of gaze cues by olive baboons (*Papio anubis*). Anim Cogn 4: 1-10.

Vick, S-J, Anderson JR (2003). The use of visual orientation cues in a competitive task by olive baboons (*Papio anubis*). J Comp Psych 117: 209–216.

Figure 1: An example of a i) congruent schematic face trial and ii) an incongruent scrambled face trial.

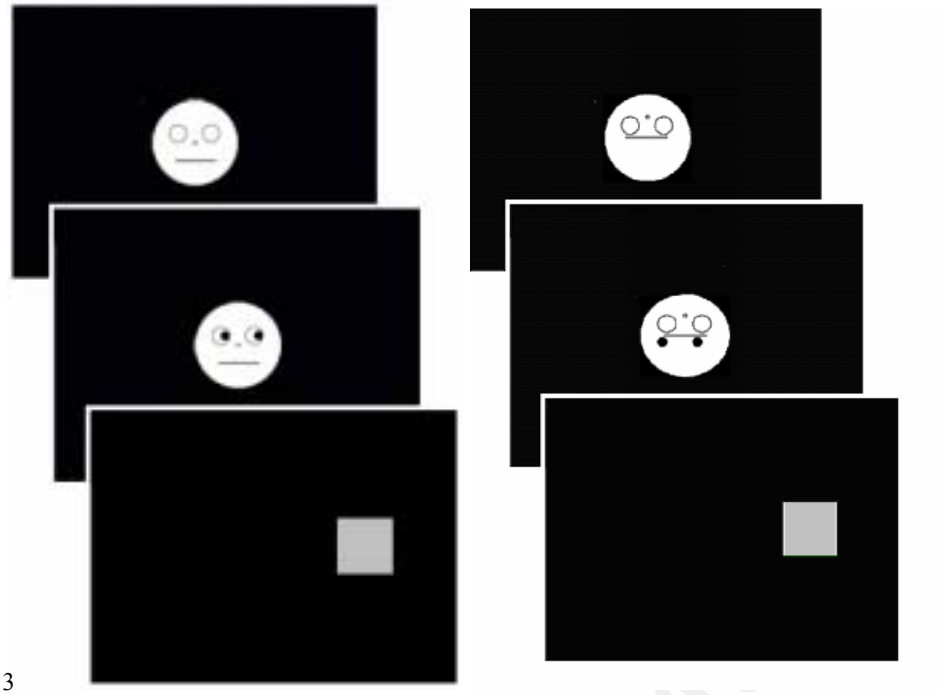

Figure 2: An example of an incongruent head trial

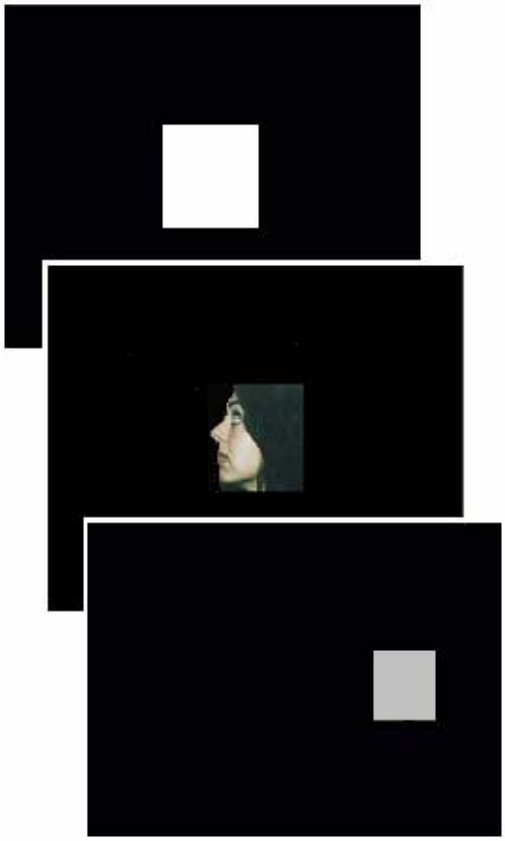

Table 1: Mean (+ S.E.) RTs (in ms) according to stimulus type, SOA and cue type.

Stimulus	SOA	Congruent	Incongruent	Neutral
Face	**100ms**	1169 (101)	1182 (57)	1206 (75)
	300ms	1329 (70)	1142 (99)	1180 (86)
	1000ms	1326 (61)	1159 (81)	1360 (84)
Control	**100ms**	1329 (66)	1312 (57)	1201 (62)
	300ms	1126 (62)	1214 (83)	1222 (100)
	1000ms	1239 (66)	1173 (72)	1240 (84)

Table 2: Mean (+S.E) RTs (in ms) for detecting targets cued by congruent and incongruent and neutral cues at 100ms, 300ms and 1000ms SOAs.

Stimulus	SOA	Congruent	Incongruent	Neutral
Head	**100ms**	959 (42)	928 (53)	--
	300ms	988 (44)	1099 (48)	--
	1000ms	1049 (67)	1028 (62)	--
Half profile	**100ms**	991 (61)	1041 (48)	--
	300ms	1012 (34)	880 (26)	--
	1000ms	1021 (34)	997 (65)	--
Eyes	**100ms**	954 (41)	916 (44)	944 (36)
	300ms	1033 (39)	1078 (32)	923 (34)
	1000ms	1010 (60)	1043 (53)	971 (43)